DOCUMENT RESUME

ED 263 516 CG 018 673

AUTHOR Blake, Catherine A.
TITLE In Response to Walker's (1984) Review on Sex Differences in Moral Development.
PUB DATE [85]
NOTE 26p.; For related document, see ED 251 749.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Developmental Stages; *Hypothesis Testing; Meta Analysis; *Moral Development; Moral Values; *Research Methodology; *Research Problems; Sex Bias; *Sex Differences; *Statistical Analysis; Value Judgment
IDENTIFIERS Kohlberg (Lawrence); *Walker (Lawrence)

ABSTRACT

Lawrence Walker's (1984) meta-analysis of 108 samples concluded that there are no sex differences in stages of moral reasoning. This paper reviews Walker's work and points out that some of his analyses, however, are controversial and favored unwarranted conclusions. Walker neither tested the hypothesis that males are more advanced than females in their moral reasoning nor did he address the allegation that Kohlberg's theory tends to underestimate the moral reasoning of women. Instead, the null hypothesis tested was that of no relationship between gender and moral reasoning. Walker's use of the Stouffer statistic and his assuming exact findings of no sex differences for many samples seem to have ensured support of the null hypothesis. Use of the counting method revealed that more samples reported gender differences in moral reasoning than would be expected if the null hypothesis were true. Observations suggest that Walker may have underestimated the extent of possible sex bias in Kohlberg's theory and scoring system. Walker's attributing his claim that his review probably overestimated the incidence of sex differences in stages of moral development to the file-drawer concept seems improper. Furthermore, Walker's vote counting analysis is biased in favor of no sex difference findings and his discursive narrative review does not seem to adhere to an explicit analytic procedure. His findings do not really indicate that there is no relationship between sex and moral reasoning, and they definitely do not suggest that further research in this area is inappropriate. (ABB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263516

In Response to Walker's (1984) Review On Sex Differences in Moral Development

Catherine A. Blake
Sexology Department, University of Quebec at Montreal

Requests for reprints should be sent to Catherine Blake, Sexology Department, University of Quebec at Montreal, P.O. Box 8888, Station "A", Montreal, Quebec, H3C 3P8, Canada.

Running Head: RESPONSE TO WALKER

CG 018673

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Catherine A. Blake
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

In Response to Walker's (1984) Review
On Sex Differences in Moral Development

Lawrence Walker (1984) undertook a laudable challenge when he synthesized conflicting findings from 108 samples. In response to an allegation of sex bias leveled against Kohlberg's theory of moral judgment development, Walker reviewed the research literature to determine whether there are consistent sex differences indicating greater moral maturity for males than for females (see Walker, 1984, p. 677). He concludes that there are no sex differences in stages of moral reasoning and recommends that we "focus our attention on other concerns" (p. 688).

The purpose of this paper is neither to prove nor to disprove the existence of sex differences in moral judgment development. Instead, it is intended to point out instances in which Walker used analyses that are controversial, if not improper, and that favour unwarranted conclusions.

Walker's Meta-Analysis

Contrary to Walker's (1984) claim, his meta-analysis does not test the hypothesis that males are more advanced than females in their moral reasoning. Nor does it address the allegation that Kohlberg's theory tends to underestimate the moral reasoning of women. Instead, the null hypothesis tested by the Stouffer procedure is that of no relationship between gender and moral reasoning.

Although the Stouffer statistic is easy to calculate, using it to summarize 108 conflicting outcomes--without first determining its suitability--undermines Walker's (1984) review. The Stouffer method requires adding the standard normal deviates (or *z*s) associated with the exact one-tailed *p* values obtained, and dividing the sum by the square root of the number of independent samples being combined. As there were no standard normal deviates and exact *p* values for many of the samples, Walker assumed exact findings of no sex difference and used $p = .5$. Therefore, the Stouffer meta-analysis involved summing up many $z = 0$ values and dividing the total by the square root of 108. Although researchers have assumed exact findings of no difference and have used $p = .5$ for some missing data, methodologists do not recommend this practice when many statistics are missing. According to Eaton (1984), the larger the ratio of missing to non-missing data, the less likely rejection of the null hypothesis becomes. Consequently, by assuming exact findings of no sex difference for many of the samples, Walker seems to have virtually ensured that the null hypothesis would be supported by default. This bias can be estimated, from the specific *z*s Walker used for his meta-analysis, and need not be presumed negligeable.

The probability of having committed a Type II error, of failing to reject the null hypothesis when it is actually false, would be further increased by heterogeneity of p or of z values (see Hunter, Schmidt, & Jackson, 1982; Rosenthal, 1984; Rosenthal & Rosnow, 1984). Although only one-tailed p values are used for calculating the overall Z value, the meta-analysis was not based on only the eight findings indicating greater moral maturity for males than for females. Instead, the Stouffer method required placing plus and minus signs before each of the z values to distinguish differences favouring males from those favouring females. Walker reports sex difference findings for 20 of the 108 samples: (a) infrequent and small differences favouring females for the Childhood and Early Adolescence samples, and (b) infrequent and small differences favouring males both for the Late Adolescence and Youth samples and for the Adulthood samples. The problem is what to do with these significant findings. With the Stouffer method, positive and negative significant outcomes in relative balance tend to cancel each other out numerically. Therefore, to lessen the probability of failing to reject a false null hypothesis, Walker should have ruled out heterogeneity of p or of z values prior to conducting the Stouffer meta-analysis.

Counting Method (χ^2)

According to Rosenthal (1984), "Practical experience with the various methods suggests that there is only rarely a serious discrepancy among appropriately chosen methods." (p. 103) Although not an elegant or powerful estimate of overall probability, the Counting method (χ^2) can be used as a simple and robust check for Walker's Stouffer finding. Had Walker have provided us with statistics for the studies he reviewed, a more powerful test could have been used.

The Counting method, using the χ^2 statistic, is a meta-analysis comparing the number of samples expected to reach a given level of significance under the null hypothesis with the number actually reaching that level.

Insert Table 1 about here

More samples reported gender differences in moral reasoning than we would expect if the null hypothesis were true, χ^2 (1, N = 108) = 8.71 , $p < .005$, or Z = 2.95 , $p < .002$, one-tailed. This finding, of a significant number of studies reporting gender differences, undermines Walker's conclusion that he has amassed support for the null hypothesis of no sex differences in moral reasoning (see Walker, 1984, p. 688). Moreover, the difference between Walker's Z value (Z = +.73) and my Z value (Z = 2.95) indicates that his meta-analysis is biased in favour of the null hypothesis. Furthermore, it is to be noted that the Z value obtained by the Counting method seems resistant to the file-drawer problem (i.e., the

common interpretation). At least 292 new, filed, or unretrieved samples averaging null results ($\underline{z} = 0$) are required to bring this overall $\underline{Z}$ value ($\underline{Z} = 2.95$) to non-significance at $\underline{p} = .5$ (see Rosenthal, 1984, p. 109).

The Allegation of Sex Bias

In his concluding remarks, Walker (1984) recommends, "Rather than arguing over the extent to which sex bias is inherent in Kohlberg's theory of moral development, it might be more appropriate to ask why the myth that males are more advanced in moral reasoning than females persists in light of so little evidence." (p. 688) What is the evidence, as provided us by Walker's review?

According to Baumrind & Pulos (1985) and to Haan (1985), only the Adulthood samples are of relevance to the charge that Kohlberg's theory underestimates the moral reasoning of women. Without taking into consideration their reanalyses of data, we find that 4 out of 21 samples indicate greater moral maturity for men (see Walker, 1984, p. 683). This finding is more than we would expect if the null hypothesis were true, χ^2 (1, $\underline{N} = 21$) = 8.72, $\underline{p} < .01$. This significant finding and the observation that none of the samples favoured the moral reasoning of women also undermine Walker's claims that there are no sex differences in stages of moral reasoning and that further research in this area would be inappropriate (see Walker, 1984, p. 688).

However, it is to be noted that much of the research in this area indeed seems inappropriate: Many studies, if not most, have based their conclusions on comparisons between Mean Moral Maturity Scores. According to Baumrind & Pulos (1985), to Blake & Cohen (1985), and to Snarey, Reimer, & Kohlberg (1985), the recommended procedures for examining sex differences are comparisons either between Global Stage Scores or between percentages of moral reasonings at each stage. Moral Maturity Scores do not preserve the necessary distinctions that would allow investigators to determine whether women follow Kohlberg's sequence of stages for moral judgment development and whether there are sex differences at stages 3, 4, and above. Consequently, the significant number of Adulthood samples, favouring the moral reasoning of men, may actually underestimate the extent of sex bias in Kohlberg's theory and scoring system.

Contrary to Walker's (1984) claim, there is some evidence for the assumption "that women fixate at Stage 3, whereas men progress to Stage 4" (p. 687). According to Parikh (1980), "The majority of the Indian mothers in the present sample are at stage 3 and the fathers at stage 4." (p. 1036) In addition, according to Holstein (1976), "Fathers . . . are modally at stage 4 or a mixture of 5 and 4 . . . , with virtually no use of stage 3 thinking" (p. 55) whereas "Many of the women are either predominantly stage 3, or if stage 4, show so much stage 3 reasoning that their stage score is a mixed one (e.g., 4[3])" (p. 60). These quotations undermine Walker's (1984) observations that "The modal

stage for both men and women in the Holstein (1976) study . . . was Stage 4, and the modal stage for both sexes in the Parikh (1980) study was Stage 3" (p. 687).

Other findings, taken from Walker's (1984) review and from studies published in Child Development[1], also suggest that there may be sex differences at stages 3, 4, and above. According to Walker (1984): (a) Levine (1976) found that women used more stage 3 reasoning than men; (b) women were overrepresented at stage 3 for the data presented by Haan, Smith, & Block (1968); and (c) Saltzstein, Diamond, & Belenky (1972) found that the moral reasoning of girls tended to be classified at stage 3, whereas that of boys tended to be classified primarily at stages 1 & 2 and secondarily at stages 4 & 5. Moreover, Holstein (1976) found that the moral reasoning of adolescent boys tended to be classified at stages 2 & 4, whereas that of adolescent girls tended to be classified primarily at stage 3 and secondarily at stage 4. Furthermore, Haan, Langer, & Kohlberg (1976) report that the moral levels of all but the youngest age group are higher for males; and Gibbs, Arnold, & Burkhart (1984) find significant sex differences in terms of stage aspects at stages 3 & 4. It is to be noted that Gibbs et al's (1984) stage 3 aspects differ from Kohlberg's (see Colby & Kohlberg, in press; Kohlberg, Levine, & Hewer, 1983).

These observations definitely suggest that Walker's (1984) review may have underestimated the extent of possible sex bias in Kohlberg's theory and scoring system. It remains to be seen whether these sex difference findings will actually be invalidated through rescoring with Kohlberg's latest manual (re., Colby & Kohlberg, in press) (see Baumrind & Pulos, 1985; Haan, 1985).

The File-Drawer Problem

Walker claims that his review probably overestimates the incidence of sex differences in stages of moral judgment development. His application of the file-drawer concept, although popular, seems improper.

When Rosenthal (1979, 1984), Glass, McGaw, & Smith (1981), and others refer to the file-drawer problem, they are referring to publication bias in favour of experimental effects. That is, studies reporting differences due either to treatment or to group effects seem more likely to get published than those reporting no differences. Thus, the typical review is likely to overestimate the extent of these differences. However, this does not mean, ipso facto, that Walker's review probably overestimates the incidence of sex differences in moral reasoning. With regard to Kohlberg's theory of moral judgment development, "Structural-developmental theory's claim to universality is not consistent with inherent sex differences in the structure of moral reasoning." (Snarey et al., 1985, p. 4) This implies that researchers may not adhere both to Kohlberg's structural-developmental theory and to a favoured

hypothesis of sex differences. Moreover, as Walker (1984) notes, "sex differences were often of secondary interest to researchers" and "many failed to report the statistics on which they made inferences of no differences" (p. 687). These quotes suggest that many of the studies, which Walker reviewed, were not intended to demonstrate sex effects. In other words, an experimental effect due to sex was not the favoured hypothesis for these studies and thus sex difference findings were unlikely to be reported. In light of this, Walker's review likely underestimates the incidence of sex differences in moral reasoning.

In summary, the common interpretation of the file-drawer problem seems both inadequate and inapplicable for Walker's review. This concept of publication bias needs to be expanded to include the notion of favoured hypothesis. When the favoured hypotheses are of universality and of no sex differences (i.e., for adherents to Kohlberg's structural-developmental theory), then publication bias is expected in the direction of no sex effects. No valid argument can be given for blindly assuming that the favoured hypothesis is always one of differences.

Walker's Vote-Counting Analysis

In addition to the Stouffer meta-analysis, Walker (1984) conducted a vote-counting (box-scoring) analysis to examine the consistency of sex differences in moral reasoning. Walker grouped the samples into one of three age categories and observed that samples reporting no sex differences were greater in number, than those reporting differences, for all three groups. Consistent with this apparent pattern, Walker (1984) concludes that "the moral reasoning of males and females is more similar than different" (p. 687). However, it is important to note that vote counts or tallies of significant findings (versus non-significant findings) neither describe nor measure the relationship between sex and moral reasoning. Therefore, Walker's conclusion seems unwarranted. Moreover, it is also to be noted that his vote-counting analysis neglects that many more sex difference findings are reported than we'd expect if the null hypothesis were true (18.5% observed versus 5% expected).

Walker's (1984) vote-counting analysis is biased in favour of no sex difference findings on three counts:

(1) Walker has ignored the fact that studies with different sample sizes have completely different meanings for significant. That is, by treating large and small samples as if they have the same power or probability of rejecting the null hypothesis when it is false, he has ignored that small samples have little power for detecting differences and thereby tend to yield non-significant results. For Walker's (1984) review, not only are many

of the samples small but (also) sample size seems to be systematically related to the significance/non-significance of the data (see Tables 1 - 3, pp. 682-686).

(2) With the vote-counting procedure, the probability of having committed a Type II error increases as the number of integrated samples increases. As Hedges & Olkin (1980) and as Hunter et al. (1982) point out, the more samples included in the review (e.g., 108 samples for Walker's review), the greater the certainty of a false conclusion about the meaning of the data. In other words, reviewers will become ever more falsely convinced that the majority of studies show no sex effect and thus that there are no sex differences in stages of moral judgment development.

(3) According to Baumrind & Pulos (1985), Walker's improper use of the Kolmogorov-Smirnov test biased the findings for some of his largest samples. These findings are said to be biased in favour of the null hypothesis.

Some, if not many, of the discrepancies in Walker's findings may represent false conflicting results: Similar differences or correlations may have been labeled significant for some samples but non-significant for others (see Glass et al., 1981; Hedges, 1984; Hedges & Olkin, 1980; Hunter et al., 1982; Light & Pillemer, 1984). For example, a similar relationship between sex and moral reasoning may exist for large samples with significant results and for small samples without. Therefore, to lessen the probability of failing to detect (and to describe) this relationship, Walker (1984) should have provided us with estimates of effect sizes for the samples

included in his review.

Walker's Discursive Narrative Review

In attempt to explain the conflicting results, Walker used a discursive narrative approach characteristic of most reviews in education, in psychology, and in sociology (see Jackson, 1978). Although this approach has been criticized by several methodologists (e.g., Glass et al., 1981; Hauser-Cram, 1984; Hunter et al., 1982; Light & Pillemer, 1982, 1984), Light & Pillemer (1982) note that "Narrative reviews are not by definition full of flaws. A careful narrative review, with explicit analytic procedures, can be extremely valuable for certain purposes." (p. 13) However, as illustrated by examples taken from the Child Development studies included in Walker's review[1], Walker does not seem to have carefully adhered to an explicit analytic procedure. Consequently, the narrative review seems especially susceptible to his particular perspective and seems an inefficient way to extract useful information (see Glass et al., 1981; Hauser-Cram, 1984; Light & Pillemer, 1982).

Further inspection of the Child Development studies reveals that Walker did not reliably adhere to his own criteria for excluding studies from the review[1]. One of Walker's (1984) criterium was to exclude studies "if there was no report or analysis of sex differences" (p. 680). By this, I assume that he means analysis of sex differences in stages of moral reasoning development. If this is what he meant, then studies focusing on sex-by-treatment interactions should have been excluded from the review (e.g., Gibbs, Widaman, & Colby, 1982;

Walker, 1982). Instead of assessing sex effects, these studies respectively assessed the validity of a questionnaire and the effectiveness of an educational intervention. By aggregating qualitative and quantitative findings from studies with different aims (i.e., analyzing sex differences in stages of moral reasoning versus analyzing sex-by-treatment and other interactions), Walker's review is unlikely to provide meaningful information about overall sex differences in moral reasoning (see Light & Pillemer, 1982).

Another of Walker's (1984) criterium was to exclude studies if subjects were selected according to stage of moral judgment development (see Walker, 1984, p. 680). This suggests that the studies done by Krebs & Gillmore (1982), by Selman (1971), and by Walker (1980, 1982) should have been excluded. Their inclusion adds to the review's selection error and weakens its internal validity. To give one example, Walker (1980) notes:

> Thus, the intervention phase involved forming three groups: (a) moral stage 2 children who had attained beginning formal operations and perspective-taking stage 3; (b) moral stage 2 children who had attained beginning formal operations and perspective-taking stage 2; and (c) moral stage 2 children who had attained concrete operations and perspective-taking stage 2. (p. 134)
>
> Children scored at other cognitive stages, including those who were transitional, were not pretested further. (p. 135)

As this study controls for stage of moral judgment

development, its inclusion biases Walker's findings in favour of no sex differences in stages of moral reasoning.

A third criterium Walker (1984) gave for excluding studies from his review was "if some objective measure of moral reasoning (such as the Defining Issues Test [DIT]) was used instead of Kohlberg's interview measure" (p. 680). As Kohlberg's interview measure was not used for several studies (e.g., Haan et al., 1968[2]; Gibbs et al., 1984; Holstein, 1976; Kahn, 1982[3]; Selman, 1971; Sullivan, McCullough, & Stager, 1970), they should have been excluded in order to minimize instrumentation error and thereby increase the internal validity of Walker's review. A fourth criterium Walker (1984) gave for excluding studies was for data previously reported in other studies (e.g., Haan, 1975[4]; Haan, Stroud, & Holstein, 1973[4]; White, Bushnell, & Regnemer, 1978[2]).

The Research Literature on Sex Differences

Although much research has been conducted in the area of moral judgment development, there seems to be an inadequate data bank with which to form conclusions about sex differences in moral orientations and development. Failure to present data for re-assessment and for comparing effect sizes, reports of conflicting results, poor experimental designs, differences in stage definitions, inconsistent usage of dilemma scoring procedures and of controls across studies, and inadequate statistical analyses are some of the intricacies to be resolved. Many studies do not test for sex effects; too many investigators assume that failure to reject the null hypothesis means that there is no relationship between

sex and moral reasoning and that there are no developmental trends in the data. Moreover, numerous researchers assume that similar Mean Moral Maturity Scores reflect more than similar overall averages, and consequently, they do not document analyses of distributions across stages. Furthermore, several investigators seem to confuse research on the existence of sex differences with research on its causes. Rather than first establishing the existence or non-existence of a sex difference in random samples of the population, as recommended by Hyde (1981), investigators proceed to analyze the causes by matching subjects for prior relevant experience (see Colby & Kohlberg, in press; Kohlberg et al., 1983; Snarey et al., 1985). Finally, these prior relevant experiences, for explaining away sex difference findings, seem to focus upon a limited number of variables, such as education, socio-economic status, and occupational choice, thereby tending to ignore other possible contributing factors, such as how people construct, resolve, and evaluate real-life moral conflicts (see Gibbs et al., 1984; Gilligan, 1982; Haan, 1985; Langdale-Hofer, 1983; Lyons, 1983; Walker, 1984).

According to McGee (1980), "A growing body of empirical evidence is beginning to substantiate that there is more to the study of behavioral sex differences than to recognize that females gestate, lactate, menstruate, and that males do not." (p. 243) There seems to be a growing awareness of sex differences in cognitive and social development (e.g., Block, 1979; Eagly & Crowley, 1985; Gold, Crombie, Brender, & Mate, 1984; Hall, 1978; Hyde, 1981; Meehan, 1984; Rosenthal & Rubin, 1982),

areas directly related to Kohlberg's theory on moral judgment development (Kohlberg, 1981; Wippman, 1982). As Kohlberg's theory concerns the development of socio-moral cognitions, research on sex differences in moral reasoning perhaps needs to be taken more seriously. In light of numerous differences in socialization experiences differentiating males from females, the existence or non-existence of sex differences in moral judgment development does not seem to have received adequate empirical investigation. More reliable and valid methods of analyses, than those traditionally used, seem required.

Conclusion

In light of the methodological weaknesses and of the criticisms raised in the present response, Walker's (1984) review should be interpreted cautiously. His findings do not really indicate that there is no relationship between sex and moral reasoning, and they definitely do not suggest that further research in this area is inappropriate.

References

Baumrind, D. (1982). Are androgynous individuals more effective persons and parents? Child Development, 53, 44-75.

Baumrind, D., & Pulos, S. (1985). Sex differences in moral judgment: Critique of Walker's conclusion that there are none. Manuscript submitted for publication.

Blake, C., & Cohen, H. (1985). A meta-analysis of sex differences in moral development. Ressources in Education, May, (MS. No. ED251749).

Block, J.H. (1979, September). Socialization influences on personality development in males and females. Paper presented at the meeting of the American Psychological Association, New York.

Colby, A., & Kohlberg, L. (Eds.). (in press). The measurement of moral judgment. New York: Cambridge University Press.

Eagly, A., & Crowley, M. (1985). Gender and helping behavior: A meta-analytic review of the social psychological literature. Manuscript submitted for publication.

Eaton, W.O. (1984). On obtaining unpublished data for research integrations. American Psychologist, 39, 1325-1326.

Gibbs, J.C., Arnold, K.D., & Burkhart, J.E. (1984). Sex differences in the expression of moral judgment. Child Development, 55, 1040-1043.

Gibbs, J.C., Widaman, K.F., & Colby, A. (1982). Construction and validation of a simplified, group-

administrable equivalent to the moral judgment interview. Child Development, 53, 895-910.

Gilligan,C. (1982). In a different voice: Psychological theory and women's development. Cambridge, Mass.: Harvard University Press.

Glass, G., McGaw, B., & Smith, M. (1981). Meta-analysis in social research. Beverly Hills: SAGE.

Gold, D., Crombie, G., Brender, W., & Mate, P. (1984). Sex differences in children's performance in problem-solving situations involving an adult model. Child Development, 55, 543-549.

Haan, N. (1975). Hypothetical and actual moral reasoning in a situation of civil disobedience. Journal of Personality and Social Psychology, 32, 255-270.

Haan, N. (1985). With regard to Walker (1984) on sex "differences" in moral reasoning. Manuscript submitted for publication.

Haan, N., Langer, J., & Kohlberg, L. (1976). Family patterns of moral reasoning. Child Development, 47, 1204-1206.

Haan, N., Smith, M.B., & Block, J. (1968). Moral reasoning of young adults: Political-social behavior, family background, and personality correlates. Journal of Personality and Social Psychology, 10, 183-201.

Haan, N., Stroud, J., & Holstein, C. (1973). Moral and ego stages in relationship to ego processes: A study of "hippies". Journal of Personality, 41, 596-612.

Hauser-Cram, P. (1984, August). Discursive and meta-

analytic reviews. Paper presented at the meeting of the American Psychological Association, Toronto.

Hedges, L.V. (1984). Research synthesis: The state of the art. International Journal on Aging and Human Development, 19, 85-93.

Hedges, L.V., & Olkin, I. (1980). Vote-counting methods in research synthesis. Psychological Bulletin, 88, 359-369.

Holstein, C.B. (1976). Irreversible, stepwise sequence in the development of moral judgment: A longitudinal study of males and females. Child Development, 47, 51-61.

Hunter, J.E., Schmidt,F.L., & Jackson, G.B. (1982). Meta-analysis. Cumulating research findings across studies. Beverly Hills: SAGE.

Hyde, J.S. (1981). How large are cognitive gender differences--a meta-analysis using omega-2 and d. American Psychologist, 36, 892-901.

Jackson, G.B. (1978). Methods for reviewing and integrating research in the social sciences. (Report No. DIS 76-20309). Washington, DC: George Washington University, Social Research Group.

Kahn, J.V. (1982). Moral reasoning in Irish children and adolescents as measured by the Defining Issues Test. Irish Journal of Psychology, 5, 96-108.

Kohlberg, L. (1981). Essays on moral development. Volume 1: The philosophy of moral development. San Francisco: Harper & Row.

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. Contributions to Human Development, 10.

Krebs, D., & Gillmore, J. (1982). The relationship among the first stages of cognitive development, role-taking abilities, and moral development. Child Development, 53, 877-886.

Langdale-Hofer, C. (1983). Moral orientation and moral development: The analysis of care and justice reasoning across different dilemmas in females and males from childhood through adulthood. Dissertation Abstracts International, 44(6), 1988-B (U. Microfilms No. 8320175).

Levine, C. (1976). Role-taking standpoint and adolescent usage of Kohlberg's conventional stages of moral reasoning. Journal of Personality and Social Psychology, 34, 41-46.

Lifton, P.D. (1985). Gender differences and moral development: Fact or fiction? Journal of Personality, in press.

Light, R.J., & Pillemer, D.B. (1982). Numbers and narrative: Combining their strengths in research reviews. Harvard Educational Review, 52, 1-26.

Light, R.J., & Pillemer, D.B. (1984). Summing up. The science of reviewing research. Cambridge, Mass.: Harvard University Press.

Lyons, N.P. (1983). Two perspectives: On self, relationships, and morality. Harvard Educational Review, 53, 125-145.

McGee, M.G. (1980). The effects of brain asymmetry on cognitive functions depends upon what ability, for which sex, at what point in development. Behavioral & Brain Research, 3, 243-244.

Meehan, A.M. (1984). A meta-analysis of sex differences in formal operational thought. Child Development, 55 1110-1124.

Parikh, B. (1980). Development of moral judgment and its relation to family environmental factors in Indian and American families. Child Development, 51, 1030-1039.

Rosenthal, R. (1979). The "file-drawer problem" and tolerance for null results. Psychological Bulletin, 86, 638-641.

Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hills: SAGE.

Rosenthal, R., & Rosnow, R.L. (1984). Essentials of behavioral research. Methods and data analysis. New York: McGraw Hill.

Rosenthal, R., & Rubin, D.B. (1982). Further meta-analytic procedures for assessing cognitive gender differences. Journal of Educational Psychology, 74, 708-712.

Saltzstein, H.D., Diamond, R.M., & Belenky, M. (1972). Moral judgment level and conformity behavior. Developmental Psychology, 7, 327-336.

Selman, R.L. (1971). The relation of role taking to the development of moral judgment in children. Child Development, 42, 79-91.

Snarey, J.R., Reimer, J., & Kohlberg, L. (1985). Development of social-moral reasoning among kibbutz

adolescents: A longitudinal cross-cultural study. Developmental Psychology, 21(1), 3-17.

Sullivan, E.V., McCullough, G., & Stager, M. (1970). A developmental study of the relationship between conceptual, ego, and moral development. Child Development, 41, 399-411.

Walker, L.J. (1980). Cognitive and perspective-taking prerequisites for moral development. Child Development, 51, 131-139.

Walker, L.J. (1982). The sequentiality of Kohlberg's stages of moral development. Child Development, 53, 1330-1336.

Walker, L.J. (1984). Sex differences in the development of moral reasoning: A critical review. Child Development, 55, 677-691.

White, C.B., Bushnell, N., & Regnemer, J.L. (1978). Moral development in Bahamian school children: A 3-year examination of Kohlberg's stages of moral development. Developmental Psychology, 14, 58-65.

Wippman, J. (1982). Divergent thinking and moral development in children. Dissertation Abstracts International, 43, 1640-B (U. Microfilms No. 8217684).

Footnotes

[1]The Child Development studies are by: Baumrind (1982); Gibbs, Arnold, & Burkhart (1984); Gibbs, Widaman, & Colby (1982); Haan, Langer, & Kohlberg (1976); Holstein (1976); Krebs & Gillmore (1982); Parikh (1980); Selman (1971); Sullivan, McCullough, & Stager (1970); and Walker (1980, 1982). The selection criterium involved maximizing reader accessibility.

[2]See Walker, 1984.

[3]See Lifton, 1985.

[4]See Haan, 1985.

Table 1
Counting Method (χ^2) for Assessing Overall Significance of Sex Differences in Moral Judgment Development

Counts	Studies reaching $p \leq .05$	Studies not reaching $p \leq .05$	$\leqq$
Obtained	20	88	108
Expected (if the null hypothesis is true)	10.8[a]	97.2[b]	108

Note.

[a]Computed from .10(N).

[b]Computed from .90(N).